



## **RAPPORT FINAL D'EXECUTION**

*Projet Hors-PO 2007-2013*

*Financé au titre de l'appel à proposition 2008*

**Etude de la biodiversité de la banane quant au métabolisme des sucres :  
Construction de la qualité organoleptique et nutritionnelle à travers la recherche  
de marqueurs moléculaires associés à la teneur en saccharose du fruit**

## SOMMAIRE

|   | Pages       |
|---|-------------|
| <b>RESUME</b>   | <b>3</b>    |
| <b>1 PRESENTATION SOMMAIRE DU PROJET (objectifs et contenu)</b>   | <b>4</b>    |
| 1.1 Situation du sujet  | 4           |
| 1.2. Objectifs spécifiques du projet  | 4           |
| <b>2 BILAN D'EXECUTION (réalisations physiques)</b>   | <b>4</b>    |
| 2.1 Choix et caractérisation biochimique des variétés utilisées   | 4           |
| 2.2 Evolution de la teneur en saccharose au cours de la maturation post récolte du fruit  | 5           |
| 2.3 Isolement des gènes codant pour les invertases pariétales et vacuolaires de banane  | 5           |
| 2.4 Expression des gènes au cours de la maturation des différentes variétés de banane   | 6-7         |
| 2.4.1 <i>Le gène MaCIN1.1/MaCIN1.2 codant pour l'ADNc MacwINV1 n'est pas l'unique gène d'invertase majoritairement exprimé au cours de la maturation</i>  | 6           |
| 2.4.2 <i>Le profil d'expression du gène MaCIN1.1/MaCIN1.2 codant pour l'ADNc MacwINV1 n'est pas toujours corrélé à l'activité invertase acide pariétale totale d'une part et à la vitesse de dégradation du saccharose d'autre part</i> | 7           |
| 2.5. Dérivation des marqueurs moléculaires  | 8-9         |
| 2.5.1 <i>Marqueurs CAPS</i>   | 8           |
| 2.5.2 <i>Marqueurs microsatellite SSR (simple sequence repeat) et identifications des SNPs (single nucleotide polymorphism)</i>   | 9           |
| <b>3. CONCLUSIONS ET PERSPECTIVES</b>   | <b>9-10</b> |
| <b>4. REFERENCES BIBLIOGRAPHIQUES</b>   | <b>10</b>   |
| <b>ANNEXES</b>  |             |

Les mécanismes physiologiques régulateurs de la teneur en saccharose du fruit ont été abordés sur un échantillon de 11 variétés de banane aux propriétés sensorielles contrastées.

L'approche de génomique fonctionnelle comparée faisant appel aux méthodes de biologie moléculaire, de biochimie et d'enzymologie a été mise en œuvre. Elle a ciblé la dégradation du saccharose, étape médiée par les invertases et préalablement identifiée comme limitante dans la régulation de la teneur en saccharose du fruit.

Dix gènes d'invertases ont été isolés et l'expression de chacun d'entre eux analysée au cours de la maturation du fruit des variétés de banane contrastées. Ces données d'expression ont ensuite été mises en corrélation avec l'évolution des activités enzymatiques d'invertase acide (pariétale et vacuolaire) d'une part, et celle des teneurs en saccharose et glucose et fructose, d'autre part.

L'ensemble des données a permis de mettre en évidence une variabilité en terme de vitesse de dégradation du saccharose et d'accumulation de glucose et fructose avec des variétés dites saccharose dominant et d'autres Gluc + Fruc dominant. Nos conditions expérimentales ont permis de mettre en évidence d'autres gènes candidats potentiels dont l'action de concert au cours de la maturation du fruit serait à l'origine de l'activité Invertase acide globale. Cependant, aucun des ces gènes n'a pu être validé comme candidat dans nos conditions expérimentales le cadre y compris le gène *MacwINV1* considéré comme tel dans nos études préliminaires.

Au moyen de l'approche bioinformatique, des marqueurs de types SSR et SNPs ont pu être dérivés à partir des séquences des gènes *MaCINs* et *MaVINs*. Pour pouvoir être utilisés comme filtres précoces au programme d'amélioration variétale, le lien l'association entre ces marqueurs et la teneur en saccharose du fruit reste cependant à établir avant utilisation au moyen d'études de génétiques poussées (d'héritabilité des caractères, recherche de QTL).

## 1. PRESENTATION SOMMAIRE DU PROJET (objectifs et contenu)

### 1.1. Situation du sujet

La qualité organoleptique et nutritionnelle est la résultante de nombreux paramètres tels que la couleur, la fermeté, la saveur (sucres et acides), les arômes, la composition minérale, en fibres, en vitamines, en molécules anti-oxydantes, ..... Ces paramètres s'établissent différemment au cours de la maturation du fruit. Ils résultent du potentiel génétique de la variété et des contraintes agro-environnementales de production, ou des conditions de récolte et de conservation. Chez nombre de fruits, la teneur en sucres solubles du type saccharose, glucose et fructose est un des paramètres déterminants qui influence leur qualité. En effet, d'une part, cette teneur globale participe à l'équilibre sucres solubles/acides organiques, dont dépend la saveur du fruit, et, d'autre part, les différents sucres solubles n'ayant pas les mêmes propriétés en terme de pouvoir sucrant ou d'index glycémique, leurs teneurs respectives contribuent également à influencer la saveur, mais aussi la qualité nutritionnelle.

Nos premières études ont suggéré que le catabolisme du saccharose, notamment l'étape régulée par l'activité enzymatique de type invertase acide codée par le gène *MacwINV1*, pourrait bien être le facteur déterminant de la régulation de la teneur du fruit en saccharose. Par conséquent, cette étape pourrait être un des points discriminants de la qualité organoleptique entre les différentes variétés de banane. Dans une démarche de contribution à la construction de la qualité par amélioration variétale par croisements, et sélection assistée par marqueurs, le gène d'invertase pariétale identifié pourrait constituer un candidat potentiel pour la définition de tels marqueurs. Cependant, sa validation comme candidat préalable à son utilisation dans un tel programme, passe par une meilleure connaissance de sa structure et de son fonctionnement, ainsi que de ses éventuels apparentés, ceci chez un échantillon de variétés plus large et génétiquement divergentes. Par ailleurs, l'activité invertase acide pouvant être à double composante (vacuolaire et pariétale), il est également nécessaire d'examiner séparément le niveau d'activité des deux formes, ainsi que le degré d'implication potentielle du (des) gène(s) vacuolaire (s) dans le contrôle de la teneur en saccharose du fruit.

### 1.2. Objectifs spécifiques du projet

Le projet INVERTASE porte sur les mécanismes de la régulation de la teneur en saccharose du fruit et l'identification des déterminismes génétiques associés. Il avait pour objectifs spécifiques :

1. d'isoler les différents membres de la famille de gènes codant pour les invertases pariétale et vacuolaire de banane.
2. de contribuer à l'identification au sein de la collection de bananiers, des variétés à cuire et dessert les plus divergentes en terme de taux d'amidon, saccharose, glucose et fructose, et les caractériser, au cours du développement et de la maturation post-récolte,
3. d'analyser l'expression des différents gènes d'invertase au cours du développement et de la maturation post-récolte de ces variétés de banane contrastée afin d'en identifier les candidats,
4. d'identifier à partir de ces gènes candidats les marqueurs moléculaires (SNPs, SSR etc..) associés à la teneur en saccharose du fruit et destinés au programme d'amélioration.

## 2. BILAN D'EXECUTION (réalisations physiques)

### 2.1. Choix et caractérisation des variétés utilisées

Le lien évident entre le métabolisme du saccharose chez les fruits et leur qualité gustative est aujourd'hui établi. Dès lors, il nous est apparu intéressant de mettre en corrélation les travaux menés sur le métabolisme du saccharose dans le cadre de ce projet et ceux réalisés au même moment sur les analyses sensorielles des différentes variétés de banane.

C'est ainsi que les 11 variétés de banane considérées comme représentatives de la variabilité sensorielle des bananes (Bugaud et al., 2011) et utilisées pour les analyses de diversité sensorielle ont également servi de matériel de départ aux travaux réalisés dans le cadre ce projet.

Les fruits récoltés à maturité commerciale ont été muris à l'air à 20°C après traitement à l'éthylène (1000 ppm/24h/20°C). Durant cette maturité post récolte, un échantillon de tissu de pulpe du fruit a été prélevé à 3, 6, 9 et 13 jours après induction de la maturation, stabilisé dans de l'azote liquide et stocké à 80°C pour des analyses ultérieures de biochimie (dosage de saccharose et d'activité enzymatique) et de biologie moléculaire (expression génique). L'ensemble des variétés utilisées dans ce projet ainsi que leurs principales caractéristiques sont présentées dans la table 3 (voir photos en annexe 2). Tenant compte du fait que ce projet s'inscrit dans un projet plus globale d'amélioration de la banane dessert d'exportation et contrairement aux objectifs de départ, le choix des variétés utilisées dans ce projet s'est essentiellement focalisé sur les variétés desserts dont quelques géniteurs diploïdes utilisés dans les schémas de croisement.

| Noms variétés           | Génotype | Caractéristiques organoleptiques présumées                                  |
|-------------------------|----------|---|
| Fraissinette (FS)       | AA       | Sucré   |
| FOU(FOU)                | ABB      | Goût miel (dessert/à cuire)   |
| Ney Pouwan (NP)         | AB       | Plantain à cuire  |
| Cavendish (CAV)         | AAA      | Sucré, standard commercial  |
| Figue pomme (FP)        | AAB      | Acidulé   |
| Figue Rose (FR)         | AAA      | Acidulé   |
| IDN110 (ID)             | AA       | Sucré, utilisé comme géniteur dans les programmes de croisement             |
| Pisang Jary Buaya (PJB) | AA       | Goût mauvais  |
| Prata Ana (PA)          | AAB      | Acidulé   |
| Pisang Lilin (PL)       | AA       | acidulé, chimique, utilisé comme géniteur dans les programmes de croisement |
| Yangambi KM5 (YK5)      | AAA      | Sucré   |

**Table 3** : Liste des variétés utilisées dans cette étude et leurs propriétés sensorielles présumées.

## 2.2. Evolution de la teneur en saccharose au cours de la maturation post récolte du fruit

Les teneurs du fruit en saccharose ainsi qu'en glucose et fructose total, produit de la dégradation du saccharose ont été mesurées au cours de la maturation. La teneur en saccharose du fruit est la résultante à la fois de son métabolisme (synthèse et dégradation) mais également de la disponibilité de son précurseur, l'amidon. D'un autre côté, les teneurs en Glucose et fructose libres du fruit peuvent provenir de la dégradation de leur précurseur immédiat (le saccharose) mais également de leur degré de mobilisation vers des voies métaboliques annexes comme la respiration, la synthèse d'amidon ou encore des polysaccharides. Afin de prendre en compte la teneur en amidon du fruit à la récolte, les teneurs en saccharose, glucose et fructose mesurées dans le fruit ont été rapportées à la quantité d'amidon de départ pour pouvoir comparer les variétés les unes des autres.

Les résultats obtenus présentent une augmentation de la teneur en saccharose durant la maturation du fruit concomitamment à celle du glucose et du fructose. Comparé à l'amidon initial, la variabilité en terme de profil et de teneur en saccharose d'une part, Gluc + fruc d'autre part, observée au sein de l'échantillon variétal ne dépend pas de la ploïdie. En effet, les ratios sucrose/amidon les plus bas (autour de 0,2) sont observés aussi bien chez les diploïdes PL (AA), NP (AB), que les triploïdes PA (AAB) et dans une moindre mesure CV (AAA), alors que les taux les plus élevés sont observés chez FRN et YK5, deux des 6 variétés triploïdes analysées. Certaines variétés peuvent être considérées comme étant saccharose dominant (FS, FRN, YK5 et ID) et d'autres Gluc + Fruc dominant (PL, CV, FOU, PA, et NP). Cette distribution est également indépendante de la ploïdie.

## 2.3. Isolement des gènes codant pour les invertases pariétales et vacuolaires de banane

A partir de la séquence du premier gène *MacwINV1* codant pour l'invertase acide d'origine pariétale disponible, trois techniques de biologie moléculaire ont été mises en œuvre simultanément :

- l'amplification in vitro d'ADN génomique de banane (Genome Walking) a permis d'obtenir deux séquences promotrices correspondant au gène *MacwINV*,
- le criblage, sous-clonage et séquençage de banque BAC pour l'isolement d'une partie des gènes apparentés à *MacwINV*
- enfin l'exploitation des premières données de séquençage du génome de la variété de bananier élite (Pahang HD).

Compte-tenu des contraintes techniques liées à leur mise en application, les deux dernières méthodes ont été réalisées en collaboration avec les chercheurs CIRAD basés à Montpellier. Ce qui nous permis de bénéficier d'un appui technique de la plate forme robotique et donc de pouvoir travailler à grande échelle, ce qui est impossible à réaliser en Guadeloupe.

La mise en commun de l'ensemble des données a permis d'identifier au total **10 séquences géniques** codant pour les invertases acides de banane dont 7 pour les formes pariétales et 3 pour les formes vacuolaires. L'analyse de leur séquence a montré que :

- i) un des gènes d'invertase pariétale le gène MaCIN5 est un pseudogène de 7188 pb dont seuls 548 codent pour une partie de la protéine d'invertase
- ii) l'ARNm *MacwINV1* correspond à deux séquences génomiques (*MaCIN1.1* et *MaCIN1.2*) identiques dans leur partie codante, mais divergentes dans leur région promotrice respective. Pour ces deux gènes, les séquences obtenues dans le cadre de l'étude se sont avérées malheureusement tronquée après le 3<sup>ème</sup> exon. Afin d'obtenir les séquences complètes de chacun des ces gènes, l'approche d'amplification d'ADN

in vitro (PCR longue distance) a été mise en œuvre au moyen des amorces spécifiques aux gènes *MaCIN1.1* et *MaCIN1.2*. Les fragments d'ADN de taille ad hoc ont été obtenus avec chacune de ces amorces. Cependant, ils n'ont pas pu être séquencés faute d'avoir pu être clonés malgré nos multiples tentatives.

- iii) L'analyse des séquences d'acides aminés déduites des séquences des gènes *MaCIN* et *MaVIN* indique que ces derniers codent pour des protéines aux caractéristiques biochimiques différentes en terme de poids moléculaire et de point isoélectrique. Ce qui laisse à penser que chacun de ces gènes correspondrait potentiellement à une isoforme enzymatique.

| Gènes            | Prom  | Ex1 | In1  | Ex2 | In2 | Ex3 | In3  | Ex4 | In4 | Ex5 | In5 | Ex6 | In6 | Ex7 | Term |
|------------------|---|-----|------|-----|-----|-----|------|-----|-----|-----|-----|-----|-----|-----|------|
| <i>MaCIN1.1*</i> | 1000  | 216 | 522  | 12  | 73  | 211 |      |     |     |     |     |     |     |     | 1000 |
| <i>MaCIN1.2*</i> | 1000  | 214 | 521  | 12  | 75  | 215 |      |     |     |     |     |     |     |     | 1000 |
| <i>MaCIN2</i>    | 1000  | 171 | 172  | 12  | 108 | 861 | 75   | 162 | 94  | 240 | 183 | 96  | 75  | 255 | 1000 |
| <i>MaCIN3</i>    | 1000  | 57  | 126  | 12  | 94  | 879 | 771  | 126 | 112 | 237 | 82  | 93  | 771 | 255 | 1000 |
| <i>MaCIN4</i>    | 1000  | 200 | 321  | 12  | 85  | 858 | 103  | 159 | 79  | 246 | 381 | 90  | 117 | 195 | 1000 |
| <i>MaCIN5**</i>  | Pseudogène de 7188 nucléotides dont seuls 548 codent pour une partie de la protéine d'invertase pariétale |     |      |     |     |     |      |     |     |     |     |     |     |     |      |
| <i>MaCIN6</i>    | 1000  | 183 | 3674 | 12  | 84  | 858 | 73   | 120 | 126 | 243 | 73  | 129 | 95  | 210 | 1000 |
| <i>MaCIN7</i>    | 1000  | 204 | 877  | 12  | 233 | 861 | 83   | 162 | 67  | 243 | 104 | 99  | 95  | 186 | 1000 |
|                  |   |     |      |     |     |     |      |     |     |     |     |     |     |     |      |
| <i>MaVIN1</i>    | 1000  | 102 | 131  | 12  | 84  | 882 | 731  | 96  | 183 | 162 | 80  | 75  | 102 | 216 | 1000 |
| <i>MaVIN2</i>    | 1000  | 57  | 138  | 12  | 94  | 868 | 1771 | 132 | 72  | 261 | 62  | 90  | 77  | 228 | 1000 |
| <i>MaVIN3</i>    | 1000  | 57  | 126  | 12  | 94  | 878 | 645  | 186 | 112 | 237 | 81  | 93  | 631 | 284 | 1000 |

**Table 1 :** Caractéristiques des séquences nucléiques correspondant aux gènes d'invertase acide d'origine pariétale (*MaCIN*) et vacuolaire (*MaVIN*) isolés chez la variété d'élite Pahang HD (diploïde, AA, 100% Homozygote) par analyse et comparaison de séquences. Les chiffres indiquent la taille (en paire de base) des différentes régions du gène dont le promoteur (Prom), les exons (Ex), les introns (In) et la partie terminale (Term).

\*Gènes tronqués après le 3<sup>ème</sup> exon. Tous les deux codent pour un même messager l'ADNc *MacwINV1* (Fils-Lycaon et al., 2011)

\*\*Speudogène

|           | <i>MaCIN1*</i> | <i>MaCIN2</i> | <i>MaCIN3</i> | <i>MaCIN4</i> | <i>MaCIN6</i> | <i>MaCIN7</i> | <i>MaVIN1</i> | <i>MaVIN2</i> | <i>MaVIN3</i> |
|-----------|----------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| <i>aa</i> | 586            | 578           | 541           | 585           | 583           | 579           | 514           | 549           | 541           |
| <i>Mw</i> | 66             | 64,8          | 59,8          | 66,8          | 66,8          | 64,9          | 56,9          | 61            | 59,8          |
| <i>pl</i> | 8,9            | 6,03          | 5,30          | 6,34          | 5,77          | 8,33          | 5,09          | 5,26          | 5,30          |

**Table 2 :** Caractéristiques dont la taille en acides aminés (aa), le poids moléculaire (mW) et le point isoélectrique (pI) des séquences polypeptidiques, correspondant aux gènes d'invertase acide d'origine pariétale (*MaCIN*) et vacuolaire (*MaVIN*) isolés chez la variété d'élite Pahang HD (diploïde, AA, 100% Homozygote) par analyse et comparaison de séquences. Les valeurs correspondant aux paramètres *Mw* exprimé en kDa et *pI* sont des valeurs théoriques déterminées au moyen du logiciel « Compute pI/MW » (Gasteiger et al., 2005). Le polypeptide correspondant au pseudogène *MaCIN5* n'est pas indiqué.

\*Les gènes *MaCIN1.1* et *MaCIN1.2* codent pour un même polypeptide *MaCIN1*.

## 2.4. Expression des gènes au cours de la maturation des différentes variétés de banane

Dans une précédente étude, nous avons identifié l'ADNc *MacwINV1* codant pour une invertase pariétale acide comme candidat potentiel au contrôle de la teneur en saccharose du fruit (Fils-Lycaon et al., 2011). Dans la présente étude, nous avons montré que :

- L'ADNc *MacwINV1* pouvait être codé par deux gènes d'invertase *MaCIN1.1* et *MaCIN1.2*
- Neuf autres gènes d'invertase d'origine pariétale (*MaCINs*) et vacuolaire et (*MaVINs*) sont apparentés à *MacwINV1*.

L'expression de l'ensemble des gènes *MaCINs* et *MaVINs* a été analysée au cours de la maturation post récolte des fruits par PCR quantitative en temps réel. Afin d'acquérir des connaissances sur le fonctionnement de ces gènes, une corrélation a été recherchée entre leur expression et l'activité totale invertase pariétale (CIN) et vacuolaire (VIN) d'une part, et les teneurs en métabolites (glucose +fructose) d'autre part. L'ensemble des données est présenté en annexe 4.

#### 2.4.1. *Le gène MaCIN1.1/MaCIN1.2 codant pour l'ADNc MacwINV1 n'est pas l'unique gène d'invertase majoritairement exprimé au cours de la maturation*

Dans nos conditions expérimentales, les ARNm de 9 des 10 gènes isolés ont été détectés au cours de la maturation des fruits chez au moins une variété. Les messagers *MaCIN5* n'ont été détectés chez aucune des variétés suggérant qu'il s'agit d'un pseudogène, non exprimé et donc probablement non fonctionnel.

L'ARNm correspondant au gène *MaCIN1.1/MaCIN1.2* est apparu comme l'unique gène le plus exprimé chez la plupart des variétés analysées excepté chez les variétés FS et PL où sont majoritairement exprimés, respectivement les gènes *MaCIN4* et *MaVIN2*. Cependant, son niveau d'expression ne dépend pas de la ploïdie.

Dans trois cas de figure, le gène *MaCIN1.1/MaCIN1.2* est apparu majoritairement exprimé en compagnie d'un ou plusieurs autres gènes (*MaVIN2* chez ID, *MaCIN3* chez YK5 et, *MaVIN2* et *MaCIN4* chez PA), lesquels constituent autant de candidats potentiels.

Le gène *MaCIN1.1/MaCIN1.2* considéré comme potentiellement candidat est transitoirement induit au cours de la maturation chez toutes les variétés y compris chez la FS où le gène est globalement exprimé à des taux faibles. La chute plus ou moins drastique des messagers *MaCIN1.1/MaCIN1.2* est observée dans les phases tardives de la maturation. Cette chute est précédée par un pic d'expression 9 jours après induction de la maturation chez PJB, 6 jours chez ID et NP, et 3 jours chez le reste de variétés. Le décalage des pics d'expression du gène *MaCIN1.1/MaCIN1.2* observé d'une variété à l'autre, peut s'expliquer par le fait que les fruits n'ont pas forcément été analysés à des stades physiologiquement équivalents. En effet, l'échantillonnage des fruits de différentes variétés sur un critère chronologique tel que nous l'avons réalisé (tous les 3 jours après induction de la maturation) n'est pas forcément synchrone à leur vitesse de maturation respective. Cependant, ce profil d'expression du gène *MaCIN1.1/MaCIN1.2* est similaire à celui obtenu dans la première étude avec l'ADNc *MacwINV1* (Fils-Lycaon et al., 2011).

Le niveau d'induction du gène varie d'une variété à l'autre et ne dépend pas de la ploïdie. En effet, dans tous les groupes d'individus analysés, de toutes les variétés analysées, PJB (AA) est celle où le gène *MaCIN1.1/MaCIN1.2* est le plus exprimé et la FS (AA) est celle où il s'exprime le moins.

#### 2.4.2. *Le profil d'expression du gène MaCIN1.1/MaCIN1.2 codant pour l'ADNc MacwINV1 n'est pas toujours corrélé à l'activité invertase acide pariétale totale d'une part et à la vitesse de dégradation du saccharose d'autre part*

Chez les variétés IDN et PL (AA), YK5 (AAA) et PA (AAB), l'augmentation durant la maturation de l'activité invertase acide totale, les parts pariétale (CIN) et vacuolaire (VIN) comprises, corrobore la vitesse d'augmentation du glucose+fructose (et la diminution du saccharose). Si l'on ne considère que le gène *MaCIN1.1/MaCIN1.2*, l'augmentation de l'activité Invertase acide totale et celle de vitesse d'évolution des métabolites sont négativement corrélées à l'expression de ce gène dont le niveau des messagers diminue dans les phases tardives. En revanche, cette corrélation devient positive si l'on prend en compte l'expression de l'ensemble des gènes d'invertase (*MaCINs* et *MaVINs*) exprimés dans le fruit. L'augmentation de l'activité invertase acide totale concomitante à la vitesse d'accumulation des métabolites serait donc probablement liée à l'action conjuguée des gènes majoritairement et séquentiellement exprimés au cours de la maturation du fruit : *MaCIN1.1/MaCIN1.2* et *MaVIN2* chez ID, PL et PA, *MaCIN1.1/MaCIN1.2* et *MaCIN3* chez YK5. Ces gènes seraient donc, chez ces variétés, des candidats potentiellement associés à la régulation de leur teneur en saccharose. Chez ces variétés, l'isoforme d'invertase correspondant au gène *MaCIN1.1/MaCIN1.2* préalablement identifié comme candidat potentiel, ne constitue donc pas à elle seule la totalité de l'activité invertase CIN et encore moins celle d'invertase acide totale quand bien même ce gène serait le plus exprimé.

Chez les variétés où le gène *MaCIN1.1/MaCIN1.2* est l'unique gène majoritairement exprimé, PJB, FRN, CV, FOU, FP et NP, la répression du gène dans les phases tardives de la maturation n'est pas toujours directement corrélée ni à l'activité invertase et ni à la vitesse d'accumulation des métabolites (sucrose, Glucose + Fructose). Le niveau d'expression du gène *MaCIN1.1/MaCIN1.2* et l'activité totale d'invertase pariétale diminuent dans les phases tardives chez les variétés NP, la CV, FRN, FP et PJB. Concomitamment, la vitesse d'accumulation du

saccharose diminue chez CV, FRN et dans une moindre mesure chez NP alors que chez la FP et PJB, cette vitesse tend à augmenter. En revanche, la diminution du saccharose n'est corrélée à l'augmentation du Gluc+Fruc que chez NP et CV. Elle diminue légèrement chez PJB alors qu'elle est constante chez la FP. Chez la FOU, la concomitance n'est observée qu'entre l'activité invertase acide totale et la vitesse d'accumulation des métabolites.

Chez FOU en revanche, la baisse du niveau d'expression du gène *MaCIN1.1/MaCIN1.2* observée dans les phases tardives s'accompagne d'une augmentation de l'activité CIN totale concomitamment à l'augmentation du Gluc+Fruc et à la diminution du Saccharose.

Enfin la variété FS, aucune corrélation n'est observée entre le profil d'expression du gène majoritairement exprimé *MaCIN4*, l'activité invertase acide et la vitesse d'accumulation des métabolites.

## 2.5. Dérivation des marqueurs moléculaires.

La disponibilité des séquences d'invertases pariétales et vacuolaires (*MaCIN* et *MaVIN*) nous a permis d'aborder la recherche de marqueurs moléculaires destinés aux études de génétique poussées, ceci indépendamment de leur caractère candidat. Durant ce projet, nous avons recherché trois types de marqueurs.

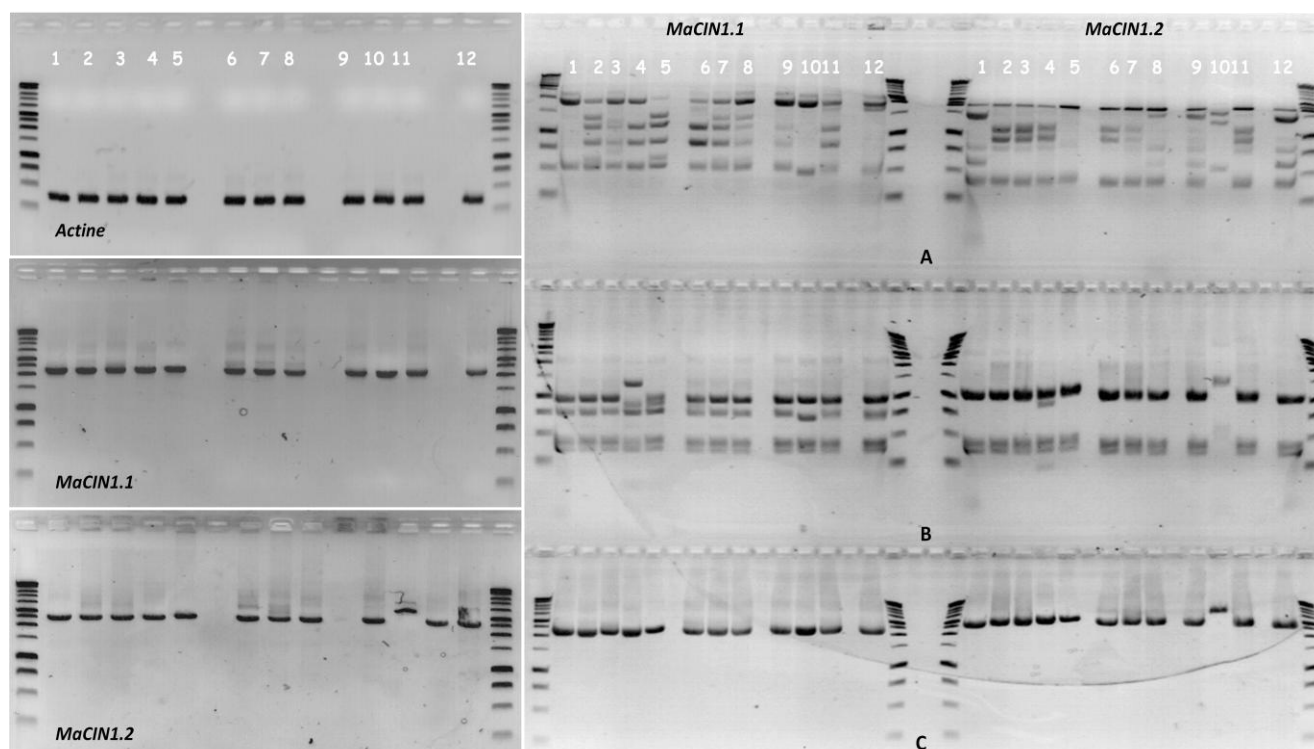
### 2.5.1. Marqueurs CAPS

Les marqueurs CAPS ont été recherchés sur les gènes *MaCIN1.1* et *MaCIN1.2*, codant pour l'ADNc MacwINV1 dont l'ARNm est le plus abondamment exprimé. Pour ce faire, l'ADN génomique de chacune des variétés a été extrait et analysé qualitativement et quantitativement. La séquence complète de chacun de ces gènes a été amplifiée séparément à partir de ces ADN au moyen des amorces spécifiques à chacun des gènes. Chaque produit d'amplification a ensuite été digéré par plusieurs enzymes *Ava I*, *Pst I* et *HIND III*. L'analyse des produits de digestion a permis de mettre en évidence une variabilité de profil de restriction (polymorphisme) d'une variété à l'autre essentiellement avec l'enzyme *Ava I*, traduisant ainsi une variabilité structurelle des gènes d'un génome à l'autre. L'ensemble amorces PCR/ Enzyme *Ava I* nous est apparu comme un marqueur potentiel en ce sens qu'il permet de discriminer les génomes porteur des gènes *MaCIN1.1* et *MaCIN1.2* (voir table ci-dessous).

| Enzymes        | Profil de restriction prédit par bioinformatique |   |
|----------------|--|---|
|                | <i>MaCIN1.1</i>                                  | <i>MaCIN1.2</i>                         |
| <i>AvaI</i>    | 1 site (753pb, 2747 pb)                          | 2 sites (573pb, 608pb, 2819pb)          |
| <i>PstI</i>    | 1 site (1105bp, 2395 pb)                         | 1 site (1530pb, 2470 pb)                |
| <i>HindIII</i> | Pas de site (3500 pb)                            | Pas de site (4000 pb)                   |
| Amorces        | CIN1.1   | CIN1.2                                  |
|                | Sens : 5'-AGCAGAAAAGCTGGAGTTGCGTAGC-3'           | Sens : 5'- TCTATACTGCCTTCGACCACAGCCG-3' |
|                | Antisens : 5'- AATTGGATCAGCACTGCGTA -3'          | Antisens : 5'- AATTGGATCAGCACTGCGTA -3' |

**Table :** Profils de restriction théorique des gènes *MaCIN1.1* et *MaCIN1.2* établis au moyen du logiciel « Restriction map » (Stothard 2000) pour les enzymes *AvaI*, *PstI* et *HindIII*. Ces profils ont été établis sur la base des séquences isolées chez le cultivar d'élite Pahang HD.





**Figure 1 : Variabilité structurelle des gènes *MaCIN1.1* et *MaCIN1.2* examinée par PCR et digestion.**

Les gènes *MaCIN1.1* et *MaCIN1.2* ont été amplifiés séparément à partir d'une quantité équivalente d'ADN génomique de différentes variétés de banane et ce au moyen des amorces spécifiques CIN1.1 et CIN1.2 (voir table ci-dessus). Après analyse des produits d'amplification sur gel d'agarose pour s'assurer qu'ils sont tous de même taille, 10µl de chaque produits sont digérés séparément par l'enzyme *Ava I* (A), *Pst I* (B) ou *Hind III* (C) et séparé sur gel d'agarose. Les différents profils obtenus permettent de révéler le marqueur CAPS (amorces CIN1.1 ou CIN1.2 et l'enzyme restriction *Ava I*) associés aux différentes tailles des fragments obtenus. Les variétés utilisées sont les suivantes, Diploïdes AA : Pahang HD (1), IDN110 (2), Pisang jary Buaya (3), Pisang mas (3), Pisang Lilin (4) ; Triploïdes AAA : Grande naine (6), Figue rose naine (7), Yangambi Km5 (8) ; autres : Figue pomme (AAB), FOU(ABB), Prata Ana (AAB). L'amplification des différents ADN au moyen des amorces d'actine (gel actine) a été réalisée pour s'assurer qu'une même quantité d'ADN a été utilisée en PCR.

### 2.5.2. Marqueurs microsatellite SSR (simple sequene repeat) et SNPs (single nucleotide polymorphisms)

Les marqueurs microsatellite ont été dérivés directement à partir des séquences des gènes *MaCIN* et *MaVIN* et au moyen du logiciel bioinformatique SAT (Dereeper et al., 2007). Au total 21 amorces microsatellites ont ainsi été identifiées (voir annexe 3).

Les marqueurs SNPs ont été définis dans les régions promotrices des gènes *MaCIN1.1* et *MaCIN1.2*. Le choix de ces régions s'explique par le fait qu'elles constituent un des sièges majeurs de la régulation de l'expression des gènes *MaCIN1.1* et *MaCIN1.2*. Par conséquent, la modification structurelle observée dans cette zone du gène peut avoir des répercussions importantes sur son expression et donc sur les voies métaboliques associées. Nous pensons donc qu'une variabilité observée à ce niveau peut être potentiellement corrélée à une variabilité de la teneur du fruit en saccharose. Les régions promotrices correspondant aux gènes *MaCIN1.1* et *MaCIN1.2* ont été amplifiées à partir de l'ADN de chacune des variétés et les fragments obtenus clonés. Après séquençage, des SNPs ont été identifiés dans les séquences dans ces régions promotrices et ce au moyen du logiciel « SNIPlay » (Dereeper et al., 2011).

## 3. CONCLUSIONS ET PERSPECTIVES

Nous avons dans cette étude, abordé les mécanismes physiologiques régulateurs de la teneur en saccharose du fruit sur un échantillon de 11 variétés de banane aux propriétés sensorielles contrastées (Bugaud et al., 2011). L'approche intégrée de biologie moléculaire, biochimie et enzymologie (génomique fonctionnelle) mise en œuvre a été ciblée sur la dégradation du saccharose médiée par les invertases, étape préalablement identifiée comme limitante dans la régulation de la teneur de ce composé (Fils-Lycaon et al., 2011).

L'ensemble des données a permis de mettre en évidence une variabilité en terme de vitesse de dégradation du saccharose et d'accumulation de glucose et fructose avec des variétés dites saccharose dominant (FS, FRN, YK5 et ID) et d'autres Gluc + Fruc dominant (PL, CV, FOU, PA, et NP). La diminution du saccharose ne s'accompagne pas toujours d'une accumulation de Gluc + fruc et inversement par conséquent, l'action des invertases n'expliquerait pas à elle seule la régulation de la teneur en saccharose du fruit.

Dans une perspective d'identification des marqueurs moléculaires, les objectifs majeurs de ce travail visaient à valider comme candidat le gène *MacwINV1* codant pour la forme pariétale de l'invertase et à identifier d'autres gènes candidats apparentés à ce gène.

Neuf membres de la famille de gènes codant pour les formes pariétales et vacuolaires d'invertase acide ont été isolés. L'approche de génomique fonctionnelle comparée telle que mise en œuvre dans cette étude n'a pas permis de valider le gène *MaCIN1.1/MaCIN1.2* comme candidat, l'hypothèse émise au terme de la première étude réalisée sur deux variétés contrastées Sowmuk « type à cuire » et IDN110 « type dessert » (Fils-Lycaon et al., 2011). Il est possible que le gène *MaCIN1.1/MaCIN1.2* soit un gène candidat plus associé à la structuration « à cuire »/ « dessert » qu'à la régulation de la teneur en saccharose du fruit.

Nos résultats ont en revanche montré que l'activité invertase acide globale n'était pas l'apanage du seul gène *MacwINV1* y compris chez les variétés où il est le plus exprimé. En effet, d'autres gènes d'invertase vacuaire (*MaVIN2* et *MaVIN3*) ou pariétale (*MaCIN3* et *MaCIN4*), prendraient une part importante, en fonction de la variété, dans l'activité invertase acide global, faisant d'eux de potentiels candidats.

Au moyen de l'approche bioinformatique, des marqueurs de types SSR et SNPs ont pu être identifiés à partir des séquences des gènes *MaCINs* et *MaVINs*. Pour pouvoir être utilisés comme filtres précoces au programme d'amélioration variétale, le lien l'association entre ces marqueurs et la teneur en saccharose du fruit reste cependant à établir avant utilisation au moyen d'études de génétiques poussées (d'héritabilité des caractères, recherche de QTL).

#### 4. REFERENCES BIBLIOGRAPHIQUES

- Bugaud C., Deverge E., Marie-Odette Daribo M.-O., Ribeyre F., Fils-Lycaon B., Didier Mbéguié-A-Mbéguié. 2011. Sensory characterisation enabled the first classification of dessert bananas. . J Sci Food Agric 91: 992–1000
- Dereeper A., Argout X., Billot C., Rami J.-F., Ruiz M. 2007. SAT, a flexible and optimized Web application for SSR marker development. BMC Bioinformatics. 8:465
- Dereeper A., Nicolas S., Lecunff L., Bacilieri R., Doligez A., Peros JP., Ruiz M., This P. 2011. SNIPlay: a web-based tool for detection, management and analysis of SNPs. Application to grapevine diversity projects. BMC Bioinformatics. 12(1):13
- Fils-Lycaon B., Julianus P., Chillet M., Galas C., Hubert O., Rinaldo D., Mbéguié-A-Mbéguié D. 2011. Acid invertase as a serious candidate to control the balance sucrose versus (glucose + fructose) of banana fruit during ripening. Scientia Horticulturae 129: 197-206
- Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A.. 2005. Protein Identification and Analysis Tools on the ExPASy Server. (In) John M. Walker (ed): The Proteomics Protocols Handbook. © Humana Press Inc., Totowa, NJ. pp 571-607. ISBN 978-1-58829-343-5.
- Stothard P (2000) The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques 28:1102-1104

## ANNEXE 1

### FICHE SYNTHETIQUE DU PROJET

#### 1. INTITULE DU PROJET

**Etude de la biodiversité de la banane quant au métabolisme des sucres : Construction de la qualité organoleptique et nutritionnelle à travers la recherche de marqueurs moléculaires associés à la teneur en saccharose du fruit.**

#### 2. COORDONNATEUR DU PROJET (nom, prénom, organisme, Tél./fax, e-mail) :

**Didier MBEGUIE-A-MBEGUIE** UMR QUALITROP, INRA-CIRAD-UAG,  
Tél. : 05 90 86 17 70 Fax : 05 90 86 80 77  
[didier.mbeguie-a-mbeguie@cirad.fr](mailto:didier.mbeguie-a-mbeguie@cirad.fr)

#### 3. RESUME DU PROJET

La filière bananière Antillaise est en crise. La banane d'exportation Antillaise est sujette à une concurrence exacerbée du fait de la libéralisation du marché et de l'uniformité de produits mis sur le marché. La disponibilité des variétés de banane résistantes aux pathogènes présents dans les Antilles et aux qualités organoleptiques et nutritionnelles accrues est indispensable et urgente pour faire face à cette crise. Elle permettrait, d'une part, de répondre aux diverses attentes des consommateurs, et, d'autre part, d'augmenter les revenus des producteurs via la segmentation du marché.

Le projet de recherche proposé s'appuie sur la diversité de variétés de banane cultivées en Guadeloupe, et porte sur le catabolisme du saccharose en glucose et fructose. Les teneurs en sucres solubles de ce type constituent, en effet, un des paramètres déterminants qui influence d'une part la qualité organoleptique à travers la saveur du fruit, et, d'autre part, la qualité nutritionnelle à travers la charge glycémique théorique.

Ce projet vise à :

- caractériser les différentes variétés de bananes cultivées en Guadeloupe en terme de teneurs en saccharose, glucose et fructose,
- développer les outils moléculaires et biochimiques impliqués dans la régulation de la teneur de ces composés dans le fruit,
- identifier, parmi ces outils, les plus importants utilisables comme marqueurs moléculaires de la qualité du fruit dans la perspective : i) de la création/sélection assistées par marqueurs de nouvelles variétés de bananes, et, ii) de la labellisation des productions antillaises.

Ces études contribueront à :

- caractériser la biodiversité, en terme de qualité, des différentes variétés de banane cultivées en Guadeloupe dans la perspective d'identifier des variétés intéressantes directement utilisables telles quel, ou comme géniteurs potentiels dans des schémas de croisement de manière à et optimiser ainsi le choix des géniteurs destinés aux programmes de croisement,
- développer une recherche d'excellence en Guadeloupe dans le domaine de la physiologie et la maturation de la banane,
- comprendre comment s'élabore la qualité de la banane, et identifier les marqueurs moléculaires, par une approche gène candidat, destinées à la sélection assistée par marqueurs

#### 4. PARTENAIRES PRINCIPAUX (laboratoires, équipes de recherche, industriels, ...)

UMR QUALITROP Guadeloupe : Equipe Physiologie du développement et de la maturation (Didier MBEGUIE-A-MBEGUIE, Bernard FILS-LYCAON, Dominique RINALDO)

#### 5. FINANCEMENT ET DUREE DU PROJET

| Budget prévisionnel k€ (TTC)   |                                  | Durée du projet              |
|--------------------------------|----------------------------------|------------------------------|
| Montant total prévisionnel     | 56,58 (hors salaires permanents) | 2 ans                        |
| Autofinancement entreprises :  |                                  | Autres ( <i>préciser</i> ) : |
| Autofinancement laboratoires : | 8,88 (15,7 %)                    | Autres ( <i>préciser</i> ) : |
| Part région :                  | 47,70 (84,3 %)                   | Autres collectivités :       |

## ANNEXE 2



**Fraissinette**



**IDN 110**



**Pisang Jary Buaya**



**Pisang Lilin**



**Figue Rose Naine**



**Figue Pomme**



**Prata Ana**

Photos des fruits des variétés utilisées dans cette étude (Bugaud et Obede, non publiés).

### ANNEXE 3

| Gènes    | Length | SSR        | Start | End  | Forward primer        | Reverse primer        | Product length |
|----------|--------|------------|-------|------|-----------------------|-----------------------|----------------|
| MaCIN1.1 | 3045   | (aat) 5    | 747   | 761  | GCCGTATATATCGAGGATTGG | GAGGAAGAGGGCACTGTTGA  | 195            |
|          | 3045   | (ctc) 7    | 1165  | 1185 | GGAACGGATCAACGGTACG   | CACTCGATTACGTGGGGAAG  | 199            |
|          | 3045   | (atg) 8    | 2626  | 2649 | Primer not found      | Primer not found      | No product     |
| MaCIN2   | 3956   | (ta) 10    | 3054  | 3073 | ATGGGAGATGAAAGCACCAG  | GTTGCTTTGGACCACGATG   | 198            |
|          | 3956   | (ac) 7     | 3073  | 3086 | ATGGGAGATGAAAGCACCAG  | GTTGCTTTGGACCACGATG   | 198            |
| MaCIN4   | 4783   | (ta) 5     | 776   | 785  | TTGCCAGTGTTTCATGCAAAT | AGAAGGTCCCAAGGGGTTTA  | 200            |
|          | 4783   | (ta) 5     | 3540  | 3549 | ACTTTTGCCGGCTTCGTT    | GGCTCCAAAACCTCTCCACAA | 213            |
|          | 4783   | (caaa) 3   | 4039  | 4050 | Primer not found      | Primer not found      | No product     |
| MaCIN6   | 7997   | (ct) 15    | 4996  | 5025 | AAAACCTGGCAAATGAGCAG  | GAGTGAGCCCAATGGATGTT  | 201            |
|          | 7997   | (ctca) 3   | 5024  | 5035 | AAAACCTGGCAAATGAGCAG  | GAGTGAGCCCAATGGATGTT  | 201            |
| MaCIN7   | 5136   | (tc) 7     | 33    | 46   | Primer not found      | Primer not found      | No product     |
|          | 5136   | (ct) 14    | 2049  | 2076 | GTCGTCTCCTCCATGATGCT  | CGGGGAACGTACGTAGTTAGT | 200            |
|          | 5136   | (ac) 11    | 2082  | 2103 | GTCGTCTCCTCCATGATGCT  | CGGGGAACGTACGTAGTTAGT | 200            |
|          | 5136   | (ct) 12    | 2173  | 2196 | CTACGTACGTTCCCGTACC   | AGTGTGCCCACACGATGTT   | 208            |
|          | 5136   | (ac) 9     | 2205  | 2222 | CTACGTACGTTCCCGTACC   | AGTGTGCCCACACGATGTT   | 208            |
|          | 5136   | (ctc) 5    | 1166  | 1180 | TGCTTCGTTGTCGACAGTA   | AAGCAACCGTGAGAGCAAAT  | 219            |
|          |        |            |       |      |                       |                       |                |
| MaVIN1   | 4797   | (ccg) 6    | 669   | 686  | TCGCAAGTTTTTCGCATATCT | GATGGTCAGGGTGGAGAAGA  | 208            |
|          | 4797   | (ctt) 5    | 1080  | 1094 | GTCCTACCCCTGGACCAACT  | TCACAACGAGTAGCCCATGA  | 201            |
| MaVIN2   | 5805   | (ga) 10    | 2683  | 2702 | TTAGCCCGCATTGTTAGAGG  | GCGAGGTGGCAGAACATTAT  | 189            |
|          | 5805   | (ga) 5     | 2831  | 2840 | TCAGGCGAAATGAAAGATCC  | TGTTTCCCCACTTGCTTCA   | 194            |
|          | 5805   | (gt) 10    | 4877  | 4896 | CCACGTCGGTTAAGATCTGG  | TTCCCGCAACAGCAAATTAT  | 206            |
|          | 5805   | (ctc) 5    | 625   | 639  | TGTTCTTCGTTTCCTCCTG   | GCTGCAAGCACTTCTTCCTC  | 182            |
|          | 5805   | (cttccc) 3 | 2131  | 2148 | ACGCCTCCAAGACGTTCTAC  | GGAACAGGACCTTCAATGCT  | 231            |
| MaVIN3   | 5256   | (ct) 8     | 1023  | 1038 | GGACCAAAGAAATGCTCCAA  | TTTGGATCTGATGCATGGAA  | 202            |
|          | 5256   | (ata) 6    | 4959  | 4976 | Primer not found      | Primer not found      | No product     |
|          | 5256   | (aaaga) 3  | 4976  | 4990 | Primer not found      | Primer not found      | No product     |

Marqueurs (couple d'amorces **Forward** et **Reverse**) microsatellites identifiés à partir des gènes *MaCIN* et *MaVIN* au moyen du logiciel SAT (Dereeper et al., 2011).

